APPLICATION FOR PATENT

5

Inventor: Emil Israel Katz

10

Title: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF
         INTEREST AND ANTIBODIES DIRECTED
         THEREAGAINST, AND METHODS, SYSTEMS AND KITS
15   FOR GENERATING AND UTILIZING EACH

FIELD AND BACKGROUND OF THE INVENTION

The present invention relates to peptides representative of polypeptides

20 of interest and antibodies directed thereagainst. More specifically, the present

invention relates to methods, systems and kits for generating and utilizing such

peptides and antibodies.

Immunoassays are the most commonly used type of diagnostic assay

and still one of the fastest growing technologies used for detection,

25 quantification and characterization of biomolecules. Although, bioassays

based phenotypic screening, receptor binding and enzymatic activity are also

commonly practiced, such bioassays cannot, in many instances, offer the same

unlimited applicability and specificity of immunoassays.

Immunoassay parameters, including, antibody specificity, cross-reactivity and labeling are continuously researched in efforts to improve the resolving power of immunoassays [Marks et al. (1992) Biotechnology 10:779-783, Soderlind et al. (1999) Immunotechnology 4:279-285, Ohlin et al.

5 (1996) Mol. Immunol. 33:47-56 and Hemminiki et al. (1998) Immunotechnology 4:59-69].

Antibody engineering

Chain shuffling: The principle of shuffling gene segments encoding individual or combinations of complementarity-determining regions (CDRs) or

10 entire variable (V) regions of an antibody is to create variation, which can be used to improve antibody affinity [Marks et al. (1992) Biotechnology 10:779-783, Soderlind et al. (1999) Immunotechnology 4:279-285 and Soderlind et al. (2000) 18:852-6]. One of the first examples was the gene shuffling of a human phage antibody specific for the hapten

15 2-phenyloxalozone. This was achieved by sequentially replacing the heavy and light chain V-region genes with repertoires of V-region genes obtained from non-immunized donors, resulting in a 300-fold increase in affinity.

Phage display: Engineered phage particles, which express a fusion protein consisting of an antibody fragment and a coat-protein can be used for

antibody selection. Such phage particles are affinity-purified on immobilized antigen and following amplification in a bacterial host, used for selecting antibodies from large repertoires (e.g., libraries) [Barbas et al. (1991) Proc. Natl. Acad. Sci. USA 88:7978-7982]. Two types of repertoires are used;

5 "immunized repertoires" and "naive repertoires" [Hoogenboom et al. (1998) Immunotechnology 4:1-20]. The main advantage of using an immunized repertoire is that antibodies selected are expected to have high affinities, while it's major drawback is the need for a new repertoire for each new antigen. In this respect, naive libraries are practical, since once constructed they can then

10 be used almost indefinitely against a diverse range of antigens. However, naive libraries have to be very large in order to yield antibodies with reasonable affinities.

Artificial antibodies: Artificial antibodies are a recent development of non-biological alternatives to antibodies, also referred to as "plastibodies"

15 [Haupt et al. (1998) Trends Biotech. 16:468-475]. The plastibody principle is based on molecular imprinting, namely, a recognition site, which is molded directly in a polymer, to thereby mimic the binding site of a natural antibody. Such polymeric constructs are used in molecularly imprinted sorbent assays (MIA) [Vlatakis et al. (1993) Nature 361:645-647], against for example

theophylline and diazepam. Plastibodies exhibit a cross-reactivity profile, towards structurally related drugs, similarly to bona fide antibodies. The major problem in the design of molecular alternatives to antibodies is that only low-affinities variants are obtained, probably due to the rigidity of the plastibody recognition site.

Other examples of molecular constructs with 'antibody-like' properties are the 'knottin scaffolds' or the Z domain of the α-helical bacterial receptor. The Z domain does not depend on intramolecular disulfide bridges, and is highly stable at extreme pH and heat conditions, thus making knottin scaffolds valuable in applications such as diagnostics and affinity purification. Although promising, the Z approach suffers from high dissociation constants and as such it is currently limited in applications.

Immunogens

Whole protein immunogens: A whole protein injection is the preferred method of obtaining high affinity antibodies capable of recognizing the antigen native form. Antibodies which are capable of recognizing the native antigen are frequently used in diagnostics, vaccination and the emerging field of antibody-based therapy. However, various difficulties are associated with whole protein immunogens such as extraction of sufficient amounts of

accurately folded protein. Furthermore the use of large protein molecules as immunogens produces antisera containing polyclonal antibodies to the numerous epitopes of the large protein molecules, although monoclonal antibody techniques utilizing whole proteins or large portions thereof as immunogens have been useful in narrowing the immunological response to such immunogens. However, a standalone monoclonal antibody technology is extremely time consuming and yields only a relatively small number of antibodies, which are capable of recognizing the immunogens specifically. Moreover, even when successful, such techniques cannot predict the biochemical identity of the antigenic epitope.

Synthetic peptide immunogens: Synthetic peptide vaccines have been actively researched over the past two decades (Arnon, 1991; Steward and Howard, 1987). However, results show that only a very small portion of monoclonal antibodies raised against short to moderate length peptides derived from the native protein recognize the synthetic peptide immunogens as well as the intact native protein [Arnheiter et al. (1981) Nature 294:278-280]. Furthermore, dissociation constants displayed by the immunocomplexes is oftentimes rather high.

Conjugating synthetic peptides to carriers has been attempted in efforts of improving the affinity of antibodies raised against such synthetic peptides to the native protein [Mariani M., et al. (1987) Mol. Immunol. 24:297-303]. For example Marianni and co-workers conjugated a synthetic peptide

5 corresponding to amino acid residues 166-174 of human chorionic somatomammotropin (hCS) amino acid sequence to elicit monoclonal antibody response to the native hCS molecule. Selected antibody clones were characterized for isotype and affinity. As expected antibodies produced against carrier conjugated hCS peptides showed, on an average, a 1000-folds

10 higher affinity towards the native hCS protein as compared to antibodies generated against non-conjugated peptides. Although promising, this approach oftentimes generates antibodies which are incapable of binding an antigen present in, or derived from, a biological sample.

Several additional approaches for developing synthetic peptide

15 immunogens are known in the art. For example, U.S. Pat. No: 6,261,569 teaches of partially or completely retro-inverso modified antigen analogues, which are capable of mimicking the immunological activity of a native peptide antigen. Although as disclosed in U.S. Pat. No: 6,261,569, these analogues induced the production of antibodies, which recognize a native peptide antigen

when administered as an immunogen to an immunocompetent host, retero inversion of peptide antigens has met with limited success in the field of vaccination. Limitations in knowledge concerning antigen-antibody binding prevent accurate predictions as to the nature and binding efficiencies of

5   antibodies elicited against an inverso, retro or retro-inverso peptide. The notion of retro inverso antigen analogues was further challenged by Lerner and co-workers (Lerner, 1984) who reported that antibodies generated against native, retro-, inverso- and retro-inverso forms of an influenza virus haemagglutinin peptide were not cross-reactive with the native peptide antigen.

10   There is thus a widely recognized need for, and it would be highly advantageous to have, peptides representative of a protein of interest, and high affinity antibodies directed thereagainst and methods of generating and using same for detecting, quantifying and/or characterizing polypeptides, such as for example, proteins contained in a biological sample.

15

SUMMARY OF THE INVENTION

According to one aspect of the present invention there is provided a method of generating a set of amino acid sequences representative of at least one polypeptide of interest, the method comprising: (a) computationally

generating a plurality of proteolytic cleavage products from the at least one polypeptide of interest; (b) computationally analyzing the plurality of proteolytic cleavage products according to at least one parameter defining a characteristic of an amino acid sequence; and (c) selecting a set of proteolytic

5   cleavage products from the plurality of proteolytic cleavage products according to predetermined criteria for each of the at least at least parameter, thereby generating the set of amino acid sequences representative of the at least one polypeptide of interest.

According to another aspect of the present invention there is provided a

10   computer readable storage media comprising a database of amino acid sequences corresponding to at least one polypeptide of interest, the database of amino acid sequences being generated by: (a) computationally generating a plurality of proteolytic cleavage products from the at least one polypeptide of interest; (b) computationally analyzing the plurality of proteolytic cleavage

15   products according to at least one parameter defining a characteristic of an amino acid sequence; and (c) storing a sequence of each of the proteolytic cleavage products thereby generating the database of amino acid sequences.

According to yet another aspect of the present invention there is provided a system for generating a database of amino acid sequences of at least

one polypeptide of interest, the system comprising a processing unit, the processing unit executing a software application configured for: (a) generating a plurality of proteolytic cleavage products from the at least one polypeptide of interest; and (b) analyzing the plurality of proteolytic cleavage products

5 according to at least one parameter defining a characteristic of an amino acid sequence.

According to still another aspect of the present invention there is provided a kit for quantifying at least one polypeptide of interest, the kit comprising a plurality of peptides being generated according to information

10 derived from computational analysis of the at least one polypeptide of interest, the computational analysis including generating a plurality of proteolytic cleavage products from the at least one polypeptide of interest.

According to further features in preferred embodiments of the invention described below, the plurality of peptides are labeled.

15 According to still further features in the described preferred embodiments the plurality of peptides are attached to a solid substrate.

According to still further features in the described preferred embodiments the plurality of peptides is contained in an individual container.

According to still further features in the described preferred embodiments the peptides are mixed in a single container.

According to still further features in the described preferred embodiments the plurality of peptides are generated via peptide synthesis or proteolytic cleavage of the at least one polypeptide of interest.

According to an additional aspect of the present invention there is provided a kit for quantifying at least one polypeptide of interest, the kit comprising a plurality of antibodies each capable of specifically recognizing at least one peptide of a plurality of peptides, the plurality of peptides being generated according to information derived from computational analysis of the at least one polypeptide of interest, the computational analysis including generating a plurality of proteolytic cleavage products from the at least one polypeptide of interest.

According to still further features in the described preferred embodiments the plurality of antibodies are labeled.

According to still further features in the described preferred embodiments the plurality of antibodies are attached to a solid substrate.

According to still further features in the described preferred embodiments the plurality of antibodies is contained in an individual container.

According to still further features in the described preferred embodiments the plurality of antibodies are mixed in a single container.

According to yet additional aspect of the present invention there is provided a method of quantifying at least one polypeptide of interest in a biological sample, the method comprising: (a) contacting the biological sample with at least one proteolytic agent, so as to obtain a proteolysed biological sample; (b) contacting the proteolysed biological sample with at least one antibody and at least one peptide of a plurality of peptides, wherein the antibody is capable of specifically binding the at least one peptide of the plurality of peptides, and further wherein the plurality of peptides are generated according to information derived from computational analysis of the at least one polypeptide of interest, the computational analysis including generating a plurality of proteolytic cleavage products from the at least one polypeptide of interest; and (c) detecting presence, absence and/or level of antibody binding to thereby quantify the at least one polypeptide of interest in the biological sample.

According to still further features in the described preferred embodiments the at least one antibody is attached to a solid substrate.

According to still further features in the described preferred embodiments the solid substrate is configured as a microarray and the at least one antibody includes a plurality of antibodies each attached to the microarray in a regio-specific manner.

According to still further features in the described preferred embodiments the at least one antibody and/or the at least one peptide is labeled and whereas step (c) is effected by quantifying the label.According to still further features in the described preferred embodiments the at least one peptide is attached to a solid substrate.

According to still further features in the described preferred embodiments the solid substrate is configured as a microarray and each of the plurality of peptides is attached to the microarray in a regio-specific manner.

According to still additional aspect of the present invention there is provided a method of generating at least one antibody specific to a polypeptide of interest, the method comprising using at least one peptide to generate the at least one antibody specific to the polypeptide of interest, wherein the at least one peptide is generated according to information derived from computational analysis of the polypeptide of interest, the computational analysis including

generating a plurality of proteolytic cleavage products from the polypeptide of interest.

According to still further features in the described preferred embodiments, computational analysis further includes analysis of the plurality

5   of proteolytic cleavage products according to at least one parameter defining a characteristic of an amino acid sequence and selection of a set of proteolytic cleavage products from the plurality of proteolytic cleavage products according to predetermined criteria for each of the at least at least parameter.

According to still further features in the described preferred

10   embodiments the plurality of proteolytic cleavage products are generated according to a proteolytic cleavage pattern of at least one proteolytic agent.

According to still further features in the described preferred embodiments the at least one proteolytic agent is selected from the group consisting of a proteolytic enzyme and a proteolytic chemical.

15   According to still further features in the described preferred embodiments the proteolytic enzyme is selected from the group consisting of trypsin, chymotrypsin, subtilisin, pepsin, V8 protease, thrombin and elastase.

According to still further features in the described preferred embodiments the proteolytic chemical is selected from the group consisting of cyanogen bromide and 2-nitro-5-thiocyanobenzoate.

According to still further features in the described preferred

5   embodiments the at least one parameter defining the characteristic of the amino acid sequence is selected from the group consisting of molecular weight, amino acid composition, hydrophobicity, hydrophilicity, charge, secondary structure, heterogeneity, length, post-translational modifications, polarity, solubility, amphipathic nature, sequence and immunogenicity.

10   The present invention successfully addresses the shortcomings of the presently known configurations by providing a novel approach for producing immunogens and antibodies directed thereto and novel immunoassays utilizing same.

15   <u>BRIEF DESCRIPTION OF THE DRAWINGS</u>

The invention is herein described, by way of example only, with reference to the accompanying drawings. With specific reference now to the drawings in detail, it is stressed that the particulars shown are by way of example and for purposes of illustrative discussion of the preferred

embodiments of the present invention only, and are presented in the cause of

providing what is believed to be the most useful and readily understood

description of the principles and conceptual aspects of the invention. In this

regard, no attempt is made to show structural details of the invention in more

5   detail than is necessary for a fundamental understanding of the invention, the

description taken with the drawings making apparent to those skilled in the art

how the several forms of the invention may be embodied in practice.

In the drawings:

FIG. 1 illustrates a system designed and configured for generating a

10   database of amino acid sequences representative of a protein of interest

according to the teachings of the present invention.

FIG. 2 illustrates a remote configuration of the system described in

Figure 1.

FIG. 3 illustrates a calibration curve of the enzyme-linked Elisa assay of

15   the present invention as obtained by adding serial dilutions of peptide probe to

a fixed quantity of antibodies recognizing the peptide probe. Data is fitted by

hyperbolic descending. The horizontal lines represent the optical density of

triplicate digested samples of P-glycoprotein expressing CHO cells (lower line)

and wild type CHO cells (upper line).

FIG. 4 illustrates the use of an antibody matrix designed and constructed according to the teachings of the present invention.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

5   The present invention is of peptides representative of polypeptides of interest and antibodies directed thereagainst. Specifically, the peptides generated according to the teachings of the present invention can be used to generate antibodies useful for detecting quantifying and characterizing proteins contained in, or derived from a biological sample.

10   The principles and operation of the present invention may be better understood with reference to the drawings and accompanying descriptions.

Before explaining at least one embodiment of the invention in detail, it is to be understood that the invention is not limited in its application to the details of construction and the arrangement of the components set forth in the

15   following description or illustrated in the drawings described in the Examples section. The invention is capable of other embodiments or of being practiced or carried out in various ways. Also, it is to be understood that the phraseology and terminology employed herein is for the purpose of description and should not be regarded as limiting.

Antibodies are widely used in the fields of diagnostics, vaccination and therapy.

To date, most approaches for generating antibodies use purified intact proteins or synthetic peptide antigens as preferred immunogens. While the former is limited by the need to extract sufficient amounts of precisely folded protein, an essentially laborious and expensive method, the latter suffers from low immunogenicity and high dissociation constants.

The present invention provides a novel approach for producing peptides representative of polypeptides of interest and antibodies directed thereagainst, which antibodies are capable of specifically recognizing an antigen while being of minimal cross-reactivity potential.

As described hereinunder and in the Examples section which follows, the present invention provides novel peptide antigens which when administered to an immunocompetent host are capable of effectively inducing the production of protein-specific antibodies which can be used to detect the protein in samples, which contain low protein concentrations and/or low exposure of protein-specific antigenic regions.

The phrase "amino acid sequences" refers herein to an oligopeptide, peptide, polypeptide, or protein sequence and fragments thereof. Such molecules can be naturally occurring or synthetic.

The phrase "polypeptide-of-interest" is used herein in reference of any naturally occurring polypeptide or antigenic fragment thereof. The polypeptide-of-interest in accordance with the present invention includes, but is not limited to pathogen derived polypeptides such as, poliomyelitis, hepatitis B, foot and mouth disease of livestock, tetanus, pertussis, HIV, cholera, malaria, influenza, rabies or diphtheria causing agents, or toxins such as robustoxin, heat labile toxin of pathogenic Escherichia coli strains, Shiga toxin from Shigella dysenteriae, polypeptides derived from multi-drug resistant strains of staphyloccocus Aureus and the like; degenerative diseases associated antigens such as the Amyloid-β-protein of Alzheimer's disease; and pregnancy and fertility associated antigens, which antigens include for example human chorionic gonadotropin and gonadotropin releasing hormone.

The polypeptide of interest can also be a tumor associated antigen such as growth factors, growth factor receptors, as well as oncogene-encoded proteins, which are expressed at significantly increased or decreased levels on tumor cells, examples include but are not limited to transferrin growth factor

(Tf), Epidermal growth factor receptor (HER1-3), Fibroblast growth factor receptor (FGFR-1), Epidermal growth factor (EGF), p21-Ras, p53, ERa, bcl-2 and the like or oncofetal antigens such as α-fetoprotein (AFP) and carcinoembryonic antigen (CEA), MAGE-1, MAGE-3, BAGE, GAGE-1,2 and the like. Alternatively the polypeptide of interest can be a tumor specific antigen, which antigens are unique to tumor cells, and do not occur on normal cells in the body. Alternatively, the polypeptide of interest can be one associated with DNA repair, or with the suppression or enhancement of apoptosis or cell senescence.

Thus, according to one aspect of the present invention there is provided a method of generating a set of amino acid sequences representative of one or more polypeptides of interest.

The method according to this aspect of the present invention is effected by several steps.

First, a plurality of proteolytic cleavage products of the at least one polypeptide of interest are computationally generated according to a known proteolytic cleavage pattern of one or more proteolytic agent.

Following computer generation, the proteolytically cleaved amino acid sequences derived from the polypeptide of interest are computationally

analyzed according to one or more parameters, which define amino acid sequence characteristics.

These parameters are used individually or in combination according to predetermined criteria to select a set of amino acid sequences derived from the at least one polypeptide of interest.

Sequences that represent the one or more polypeptides of interest can be stored in a database which can be generated by a suitable computing platform. Amino acid sequences of the present invention can be further used to generate peptides and antibodies directed thereagainst, which can be packed in diagnostic kits and implemented in various therapeutic and diagnostic methods.

The proteolytic cleavage products described above can be computationally generated according to a known proteolytic cleavage pattern of one or more proteolytic agents. This can be effected using any processing software capable of recognizing proteolytic cleavage sites within an amino acid sequence and generating the amino acid sequences of such proteolysis.

It will be appreciated that although the use of a dedicated software application, such as Sciprot (available from www.asiaonline.net.hk/~twcbio/DOCS/1/scPrtein.htm), the GCG package (Genetics Computer Group, Wisconsin) or Macvector (available from

www.accelrys.com/products/macvector/) is preferred, computational

proteolysis of amino acid sequences can also be generated using non-dedicated

software applications such as for example, a text edit application (e.g., Word

available from Microsoft Inc.). Such a non-dedicated application can be used

5   to recognize specific alphanumeric character sequences (i.e., cleavage sites)

and to generate sequences of amino acids, which correspond to the cleavage

products according to commands provided by the user (e.g., a Word Macro).

The proteolytic agent can be any agent, which is capable of cleaving

polypeptides between specific amino acid residues (i.e., the proteolytic

10   cleavage pattern).

According to one embodiment of this aspect of the present invention a

proteolytic agent is a proteolytic enzyme. Examples of proteolytic enzymes,

include but are not limited to trypsin, chymotrypsin, V8 protease, pepsin,

subtilisin, thrombin, elastase, caspase-1, caspase-2, caspase-3, caspase-4,

15   caspase-5, caspase-6, caspase-7, caspase-8, MetAP-2, adenovirus protease,

HIV protease and the like.

According to another embodiment of this aspect of the present invention

a proteolytic agent is a proteolytic chemical such as cyanogen bromide and

2-nitro-5-thiocyanobenzoate.

It will be appreciated that although any proteolytic agent known in the art can be used by the present invention, proteolytic reagents or combinations thereof, which enable complete substrate proteolysis are preferred.

Following computer generation, the proteolytically cleaved amino acid sequences derived from the polypeptide of interest are computationally analyzed according to one or more parameters, which define amino acid sequence characteristics.

Examples of such parameters include but are not limited to molecular weight, amino acid composition, hydrophobicity, hydrophilicity, charge, secondary structure, heterogeneity, length, post-translational modifications, polarity, solubility, amphipathic nature, sequence and immunogenicity.

Each of the peptide products computationally generated is classified according to one or preferably several parameters, which are described in detail below.

Each parameter can be considered separately according to predetermined criteria or in combination with other parameters used, in which case, each parameters is also weighted according to its importance. In any case, each of the peptide products is scored. The score of each peptide may be an absolute score, in which case peptides, which score above a predetermined

threshold are selected, or alternatively, the score can be proportional in which case highest scoring peptides are selected.

Although sequence analyses can be effected using various protein analysis softwares, analysis may also be effected at least in part, using other software applications not dedicated for such use. Examples of non-dedicated software applications which can be used include spread-sheet software, such as Microsoft Excel (available from Microsoft incorporation).

The following describes in detail parameters, which can be used by the present invention to qualify peptide sequences.

Homology level: The sequences of the computer generated peptide products are compared with protein databases in order to identify substantially non-homologous, (protein-unique) sequences. Preferably, sequences are compared to all known protein databases, employing a sequence alignment algorithm such as BLAST (Basic Local Alignment Search Tool, available through www.ncbi.nlm.nih.gov/BLAST) or the Smith-Waterman algorithm. Sequences that display low homology to database sequences, e.g., not exceeding 40%, preferably not exceeding 30%, more preferably not exceeding 20%, most preferably not exceeding 10%, or presently preferred displaying

0-5% homology are selected or given a high score depending on the type and number of parameters used for qualification.

Immunogenicity: The computer generated peptide products can also be analyzed for their ability to induce a specific immunogenic response, specifically, an antibody response. Various sequence analysis softwares are known in the art, which provide an immunogenicity index according to, for example, the Jameson-Wolf algorithm. Examples include, but are not limited to, Sciprot (available from www.asiaonline.net.hk/~twcbio/ DOCS/1/scPrtein.htm) and Macvector (available from www.accelrys.com/ products/macvector/) as well as the widely utilized GCG package (Genetics Computer Group, Wisconsin).

Immunogenicity is determined, at least in part, by the following properties of the immunogen:

Foreignness- in order to elicit an immune response, amino acid sequences must be recognized as non-self by the immune system of the immunocompetent host. Generally, the greater the phylogenetic distance, between the two species (the species from which the antigen was extracted, and the immunocompetent host species), the greater the genetic and therefore the antigenic disparity between them.

Molecular size- there is a correlation between the size of an amino acid sequence and its immunogenicity. As such, amino acid sequences having a molecular mass of at least1000 daltons (Da), more preferably at least 5000 Da, even more preferably at least 10,000 Da and most preferably have a molecular mass approaching 100,000 Da are favored by the present invention and as such are either selected or given a high score.

Chemical composition and heterogeneity- In general, homopolymers (i.e., polymers composed of a single amino acid) tend to lack immunogenicity, regardless of their size. Copolymers of sufficient size, containing two or more different amino acids, are immunogenic. Furthermore, all four levels of protein organization-primary, secondary, tertiary and quaternary-contribute to the structural complexity of a polypeptide and hence affect its immunogenicity.

Susceptibility to antigen processing and presentation- Basically, polypeptides that cannot be degraded and presented with MHC molecules are poor immunogens. Moreover, large insoluble molecules are more immunogenic than small soluble ones, because they are more readily phagocytosed and processed.

Post-translational modifications: The computer generated peptide products can also be analyzed according to the presence, absence and/or level

of post-translational modifications. More than 100 different such modifications of amino acid residues are known, examples include but are not limited to phosphorylation, acetylation, methylation, hydroxylation, carboxylation and glycosylation. Sequence analysis softwares which are capable of determining putative post-translational modification in a given amino acid sequence include the NetPhos server which produces neural network predictions for serine, threonine and tyrosine phosphorylation sites in eukaryotic proteins (available through http://www.cbs.dtu.dk/services/ NetPhos/), GPI Modification Site Prediction (available through http://mendel.imp.univie.ac.at/gpi) and the ExPASy proteomics server for total protein analysis (available through www.expasy.ch/tools/)

Generally, preferred peptide products are those lacking any post-translational modification sites, since post-translationally modified amino acid sequences are often difficult to purify, and are frequently poor immunogens.

Notwithstanding from the above, peptide products which include post-translational modification, which indicate a biological activity of the polypeptide-of-interest can also be used by the present invention. A very common example is the phosphorylation of OH group of the amino acid side

chain of a serine, a threonine, or a tyrosine group in a polypeptide. Depending on the polypeptide, this modification can increase or decrease its functional activity.

Size - Preferred amino acid sequences include fragments of a molecular weight between 600 - 1800 Da and amino acid sequence length between 5 - 12. The size limitation is mainly due to the requirement for minimizing homologies to other polypeptides and technical difficulties associated with synthesizing, purifying and folding large polypeptides (besides the immunogenic-associated size limitation, discussed hereinabove).

The parameters described above are then used individually or in combination to analyze the computationally generated peptide products and to select a set of peptides most suitable for use with the present invention.

As mentioned hereinabove, selection can be effected on the basis of a single parameter or several parameters considered individually or in combination.

In cases where several parameters are examined, a scoring system e.g., a scoring matrix, is preferably used.

Since in some cases immunogenicity may be more important than both post-translational modifications and sizes, while in others, sequence homology

might be the most significant parameter, the use of a scoring matrix in which each parameter is weighted enables one to select the most suitable peptides.

Such a scoring matrix can list the various amino acid sequences (of the peptide products) across the X-axis of the matrix while each parameter can be

5   listed on the Y-axis of the matrix. Parameters include both a predetermined range of values from which a single value is selected from each peptide, and a weight. Each peptide is scored at each parameter according to its value and the weight of the parameter.

Finally, the scores of each parameter of a specific peptide are summed

10   and the results are analyzed.

Peptides which exhibit a total score greater than a particular stringency threshold are grouped as members of a peptide set; the higher the score the more stringent the criteria of grouping.

Alternatively, a set of peptides exhibiting the highest scores can also be

15   selected.

The sequences of these peptides, which represent a polypeptide of interest, can then be used to generate a database which can be stored on a computer readable media such as a magnetic, optico-magnetic or optical disk.

In addition to sequence information, such a database can also include additional data relating to database generation, parameters used for selecting peptide sequences, putative uses of the stored sequences, and various other annotations and references which relate to the stored sequences or polypeptide

5   from which they were generated

According to another aspect of the present invention and as illustrated in Figure 1, the database of peptide sequences of the present invention is generated by a system designed and configured for such function, which system is referred to hereinunder as system **10**.

10   System **10** includes a processing unit **12**, which executes a software application designed and configured for generating and preferably analyzing the plurality of proteolytic cleavage products from one or more polypeptides as described hereinabove. System **10** may also include a user input interface **14** (e.g., a keyboard and/or a mouse) for inputting database or database related

15   information, and a user output interface **16** (e.g., a monitor) for providing database information to a user.

System **10** of the present invention may be used by a user to query stored sequences, to retrieve peptide sequences stored therein or to generate peptide sequences from user inputted sequences.

System **10** can be any computing platform known in the art including but not limited to, a personal computer, a work station, a mainframe and the like.

The database generated and stored by system **10** can be accessed by an

5   on-site user of system **10**, or by a remote user communicating with system **10**.

Figure 2 illustrates a remote configuration of system **10** of the present invention.

In such a configuration, the communication between a remote user **18** and processing unit **12** is effected through a communication network **20**.

10   Communication network **20** can be any private or public communication network including, but not limited to, a standard or cellular telephony network, a computer network such as the Internet or intranet, a satellite network or any combination thereof.

As illustrated in Figure 2, communication network **20** includes one or

15   more communication servers **22** (one shown in Figure 2) which serves for communicating data pertaining to the polypeptide of interest between remote user **18** and processing unit **12**.

It will be appreciated that existing computer networks such as the Internet can provide the communication and applications necessary for

supporting data communication between any number of sites **24** and remote analysis sites **26**.

For example, using a computer operating a Web browser application and the World Wide Web, any polypeptide of interest can be "uploaded" by user **18** onto a Web site maintained by a database server **28**. Following uploading, database server **28** which serves as processing unit **12** can be instructed by the user to processes the polypeptides as is described hereinabove.

Following such processing, which can be performed in real time, peptide sequence results can be displayed at the web site maintained by database server **28** and/or communicated back to site **24**, via for example, e-mail communication.

Thus, using the Internet, a remote configuration of system **10** can provide protein analysis services to a plurality of sites **24** (one shown in Figure 2). It will be appreciated that this configuration of system **10** of the present invention is especially advantageous in cases where polypeptide analysis can not be effected on-site. For example, laboratories which lack the equipment necessary for executing the analysis or lack the necessary skills to operate it.

The peptide set selected according to the teachings of the present invention can be used to generate peptides either through enzymatic cleavage

of the protein from which they were generated and selection of peptides, or preferably through peptide synthesis methods.

Proteolytically cleaved peptides can be separated by chromatographic or electrophoretic procedures and purified and renatured via well known prior art methods.

Synthetic peptides can be prepared by classical methods known in the art, for example, by using standard solid phase techniques. The standard methods include exclusive solid phase synthesis, partial solid phase synthesis methods, fragment condensation, classical solution synthesis, and even by recombinant DNA technology. See, e.g., Merrifield, J. Am. Chem. Soc., 85:2149 (1963), incorporated herein by reference. Solid phase peptide synthesis procedures are well known in the art and further described by John Morrow Stewart and Janis Dillaha Young, Solid Phase Peptide Syntheses (2nd Ed., Pierce Chemical Company, 1984).

Synthetic peptides can be purified by preparative high performance liquid chromatography [Creighton T. (1983) Proteins, structures and molecular principles. WH Freeman and Co. N.Y.] and the composition of which can be confirmed via amino acid sequencing.

Due to their protein specificity and immunogenicity, peptides produced according to the teachings of the present invention can be used to generate antibodies characterized by high affinity and specificity.

The peptides generated according to the teachings of the present invention or the antibodies directed thereagainst can be used for both diagnostic and therapeutic purposes.

For example, peptides corresponding to selected amino acid sequences of a protein of interest can be directly administered to an immunocompetent host as immunogens in order to elicit efficient production of antibodies directed at such a protein of interest. Such antibodies would be characterized by high affinity binding and specificity and as such, in cases of disease related protein, such peptides can be used as efficient therapeutic agents.

Alternatively, such peptides can be used to generate antibodies (monoclonal or polyclonal), which in turn can be used for diagnostic purposes.

Various hosts including goats, rabbits, rats, mice, humans and others, may be immunized by peptide injection for the purposes of generating antibodies. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to Freund's mineral gels such as aluminum hydroxide, and surface active

substances such as lysolecithin, pluronic polyols, polyanions, oil emulsions, keyhole limpet hemocyanin and dinitrophenol.

Monoclonal antibodies may be prepared using any technique which produces antibody molecules by continuous cell lines in culture. These include but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the Epstein-Bar-Virus (EBV)-hybridoma technique [Kohler G., et al. (1975) Nature 256:495-497, Kozbor D., et al. (1985) J. Immunol. Methods 81:31-42, Cote R.J. et al. (1983) Proc. Natl. Acad. Sci. 80:2026-2030, Cole S.P. et al. (1984) Mol. Cell. Biol. 62:109-120].

In addition, techniques developed for the production of "chimeric antibodies", the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity can be used [Morison S.L. et al. (1984) Proc. Natl. Acad. Sci. 81:6851-6855, Neuberger M.S. et al. (1984) Nature 312:604-608, Takeda S. et al. (1985) Nature 314:452-454]. Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art.

Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening immunoglobulin libraries or panels of

highly specific binding reagents as disclosed [Orlandi D.R. et al. (1989) Proc.

Natl. Acad. Sci. 86:3833-3837, Winter G. et al. (1991) Nature 349:293-299].

Antibody fragments may also be generated. For example, such

fragments include F(ab')2 fragments which may be produced by pepsin

5   digestion of the antibody molecule and the Fab fragments which can be

generated by reducing the disulfide bridges of the F(ab')2 fragments.

Alternatively, Fab expression libraries may be constructed to allow rapid and

easy identification of monoclonal Fab fragments with the desired specificity

[Huse W.D. et al. (1989) Science 254:1275-1281].

10   Various immunoassays may be used for screening antibodies having the

desired specificity. Numerous protocols for competitive binding or

immunoradiometric assays using either polyclonal or monoclonal antibodies

with established specificity are well known in the art. Such immunoassays

typically involve the measurement of complex formation between the

15   immunogenic peptide and its specific antibody. Thus the peptides of the

present invention may be used for identifying and purifying antibodies.

It will be appreciated that although generation of antibodies against

synthetic peptides or cleavage products is preferred, antibodies generated

against the whole protein and affinity purified against the synthetic peptides or cleavage products can also be used by the present invention.

Peptides generated according to the teachings of the present invention or antibodies specific thereto can be used to diagnose a variety of diseases

5   including but not limited to diabetes, Parkinson, Alzheimer' disease, HIV, malaria, cholera, influenza, rabies, diphtheria, breast cancer, colon cancer, cervical cancer, melanoma, lung cancer, ovarian cancer, pancreatic cancer, prostate cancer, lymphomas, leukemias and the like and any other diseases which are associated with aberrant expression of multiple antigens.

10   While most current immuno-diagnostic methods are rate limited by the number of antibodies which can be applied on a given biopsy and sensitivity-limited by masking of protein-specific antigenic regions, the present invention provides an immuno-detection assay which circumvents these limitations.

15   Thus, according to an additional aspect of the present invention there is provided a method of quantifying at least one polypeptide of interest in a biological sample.

The method includes several method steps, schematically illustrated in Figure 4.

In the first step, a biological sample is obtained. The biological sample as used herein refers to any body sample such as blood (serum or plasma), sputum, ascites fluids, pleural effusions, urine, biopsy specimens, isolated cells and/or cell membrane preparation. Methods of obtaining tissue biopsies and

5   body fluids from mammals are well known in the art.

Retrieved biological samples can be further solubilized using detergent-based or detergent free (i.e., sonication) methods, depending on the biological specimen and the nature of the examined polypeptide (i.e., secreted, membrane anchored or intracellular soluble polypeptide). Oftentimes, protease

10   and phosphatase inhibitors are included within the solubilization buffer, to avoid non-regulated endogenous protease activity, and maintenance of the polypeptides active form.

The solubilized biological sample is contacted with one or more proteolytic agents. Digestion is effected under effective conditions and for a

15   period of time sufficient to ensure complete digestion of the diagnosed polypeptide(s). Agents that are capable of digesting a biological sample under moderate conditions in terms of temperature and buffer stringency are preferred. Measures are taken not to allow non-specific sample digestion, thus the quantity of the digesting agent, reaction mixture conditions (i.e., salinity

and acidity), digestion time and temperature are carefully selected. At the end

of incubation time proteolytic activity is terminated to avoid non-specific

proteolytic activity, which may evolve from elongated digestion period, and to

avoid further proteolysis of other peptide-based molecules (i.e., purified

5   peptides and/or antibodies), which are added to the mixture in following steps.

In the next method step the proteolysed biological sample is contacted

with one or more antibodies, which are capable of binding one or more

proteolytic cleavage products of the examined polypeptide(s) of interest. Such

antibodies are capable of specifically binding peptides representative of the

10   polypeptide(s) of interest, which were generated as described hereinabove.

The antibodies are attached to a solid substrate, which may consist of a

particulate solid phase such as agarose, sepharose or sephadex beads or a solid

substrate configured as an antibody microarray, such as a 96 well plate (see

Examples 4 and 5 of the Examples section which follows).

15   Contacting the proteolysed biological sample with one or more

antibodies is effected under conditions suitable for the formation of immune

complexes (primary immune complexes).

Immunocomplexes are washed to remove any non-specifically bound antibody species, allowing only those antibodies specifically bound within the primary immune complexes to be detected.

In general monitoring of immunocomplex formation is well known in the art and may be achieved by any one of several approaches. These approaches are generally based on the detection of a label or marker, such as any radioactive, fluorescent, biological or enzymatic tags or labels of standard use in the art. U.S. Patents concerning the use of such labels include U.S. Pat. No. 3,817,837; 3,850,752; 3,939,350; 3,996,345; 4,277,437; 4,275,149 and 4,366,241, each incorporated herein by reference.

The examined polypeptide(s) may be linked to a detectable label, to allow simple detection of this labeled polypeptide, thereby allowing the amount of the primary immune complexes in the composition to be determined. Polypeptide labeling can be effected by labeling the biological sample, prior to, concomitant with or following sample digestion (see Example 3 of the Examples section).

Alternatively, labeled synthetic peptides added to the reaction mixture can be used to quantify binding of the proteolytic products via competitive binding approaches.

The intensity of signal produced in any of the detection methods described hereinabove may be analyzed manually or using a computer program.

In general, polypeptide quantification is preferably effected alongside a calibration curve so as to enable accurate protein determination (see Example 4 of the Examples section below for further detail). Furthermore, quantifying polypeptide(s) originating from a biological sample of a patient is preferably effected by comparison to a normal sample, which sample is characterized by normal expression pattern of the examined polypeptide(s).

It will be appreciated that the detection method described above can also be effected using peptide rather than antibody arrays. In such a case, peptides are attached to a solid support and used along with corresponding antibodies and proteolysed biological samples in competitive binding assays aimed at detecting the presence, absence and/or quantity of specific polypeptides.

It will further be appreciated that in cases of polypeptides which are associated with the formation of anti-self antibodies, such as the case with autoimmune disease associated polypeptides, such peptide arrays can also be

used to detect the presence of such autoantibodies, thereby enabling the detection of the disease.

Immunodetection methods of the present invention have evident utility in the diagnosis of various diseases and conditions. In addition, such methods can also be used in non-clinical applications, such as, for example, antigen titering and the like.

The peptides or antibodies generated according to the present invention can be included in a diagnostic or therapeutic kit. For example, peptide sets of specific disease related proteins or antibody populations directed thereagainst can be packaged in a one or more containers with appropriate buffers and preservatives and used for diagnosis or for directing therapeutic treatment. Thus, the peptides or antibodies can be each mixed in a single container or placed in individual containers. Preferably, the containers include a label. Suitable containers include, for example, bottles, vials, syringes, and test tubes. The containers may be formed from a variety of materials such as glass or plastic.

In addition, other additives such as stabilizers, buffers, blockers and the like may also be added.

The peptides or antibodies of such kits can also be attached to a solid support, such as beads, array substrate (e.g., chips) and the like and used for diagnostic purposes. Example 4 and Example 5 of the Examples section describe the use of substrate immobilized antibody arrays designed for such purposes.

5   Peptides and antibodies included in kits or immobilized to substrates may be conjugated to a detectable label such as an enzyme, in which case the kit also includes substrates and cofactors required by the enzyme to produce a colorimetric reaction (e.g. a substrate precursor which provides the detectable chromophore or fluorophore). Alternatively, the detectable label can be a tag

10   such as an epitope tag , examples of which include but are not limited to a Myc tag, a Flag tag, a His tag, a Leucine tag, an IgG tag, a streptavidin tag and the like, in which case the kit will include an antibody directed at the epitope and a secondary labeled antibody conjugated to a chromophore or a fluorophore, possibly the epitope directed antibody is labeled.

15   The kit can also include instructions for determining if the tested subject is suffering from, or is at risk of developing, a condition, disorder, or disease associated with expression of the polypeptide of interest.

The peptides and antibodies directed thereagainst of the present invention are valuable to the fields of biomolecule research, therapy and

diagnostics. The ability to simultaneously identify multiple antigens associated

with a disease or condition can result in an optimized treatment regimen as

well as enable identification of an onset or an early stage of diseases, thereby

significantly improving prognosis and treatment.

5   Additional objects, advantages, and novel features of the present

invention will become apparent to one ordinarily skilled in the art upon

examination of the following examples, which are not intended to be limiting.

Additionally, each of the various embodiments and aspects of the present

invention as delineated hereinabove and as claimed in the claims section below

10   finds experimental support in the following examples.

EXAMPLES

Reference is now made to the following examples, which together with

the above descriptions, illustrate the invention in a non limiting fashion.

Generally, the nomenclature used herein and the laboratory procedures utilized

15   in the present invention include molecular, biochemical, microbiological and

recombinant DNA techniques. Such techniques are thoroughly explained in

the literature. See, for example, "Molecular Cloning: A laboratory Manual"

Sambrook et al., (1989); "Current Protocols in Molecular Biology" Volumes

I-III Ausubel, R. M., ed. (1994); Ausubel et al., "Current Protocols in

Molecular Biology", John Wiley and Sons, Baltimore, Maryland (1989);

Perbal, "A Practical Guide to Molecular Cloning", John Wiley & Sons, New

York (1988); Watson et al., "Recombinant DNA", Scientific American Books,

New York; Birren et al. (eds) "Genome Analysis: A Laboratory Manual

5   Series", Vols. 1-4, Cold Spring Harbor Laboratory Press, New York (1998);

methodologies as set forth in U.S. Pat. Nos. 4,666,828; 4,683,202; 4,801,531;

5,192,659 and 5,272,057; "Cell Biology: A Laboratory Handbook", Volumes

I-III Cellis, J. E., ed. (1994); "Current Protocols in Immunology" Volumes I-III

Coligan J. E., ed. (1994); Stites et al. (eds), "Basic and Clinical Immunology"

10   (8th Edition), Appleton & Lange, Norwalk, CT (1994); Mishell and Shiigi

(eds), "Selected Methods in Cellular Immunology", W. H. Freeman and Co.,

New York (1980); available immunoassays are extensively described in the

patent and scientific literature, see, for example, U.S. Pat. Nos. 3,791,932;

3,839,153; 3,850,752; 3,850,578; 3,853,987; 3,867,517; 3,879,262; 3,901,654;

15   3,935,074; 3,984,533; 3,996,345; 4,034,074; 4,098,876; 4,879,219; 5,011,771

and 5,281,521; "Oligonucleotide Synthesis" Gait, M. J., ed. (1984); "Nucleic

Acid Hybridization" Hames, B. D., and Higgins S. J., eds. (1985);

"Transcription and Translation" Hames, B. D., and Higgins S. J., eds. (1984);

"Animal Cell Culture" Freshney, R. I., ed. (1986); "Immobilized Cells and

Enzymes" IRL Press, (1986); "A Practical Guide to Molecular Cloning" Perbal, B., (1984) and "Methods in Enzymology" Vol. 1-317, Academic Press; "PCR Protocols: A Guide To Methods And Applications", Academic Press, San Diego, CA (1990); Marshak et al., "Strategies for Protein Purification and

5   Characterization - A Laboratory Course Manual" CSHL Press (1996); all of which are incorporated by reference as if fully set forth herein. Other general references are provided throughout this document. The procedures therein are believed to be well known in the art and are provided for the convenience of the reader. All the information contained therein is incorporated herein by

10   reference.

Background

Multi-drug resistance (MDR) represents a major obstacle in the successful therapy of neoplastic diseases. Studies have demonstrated that this form of drug resistance occurs both in cultured tumor cell lines as well as in

15   human cancers. Recent findings show that numerous molecular mechanisms operate in MDR; from decrease in drug cellular accumulation to the abrogation of drug-induced apoptosis. The most investigated mechanisms for MDR include activation or upregulation of transmembrane proteins, effluxing different chemical substances from the cells (P-glycoprotein is the most

characterized efflux pump), activation of the glutathione detoxification system and alterations of genes and proteins involved in the control of apoptosis (especially p53 and Bcl-2).

It has been suggested that expression of MDR associated proteins has a

5   prognostic value to various types of human cancers including leukemias and soft tissue sarcomas. For example, the failure to achieve complete remission in acute myeloid leukemia has been associated with P-glycoprotein expression [Broxterman HJ. et al. (1997) J. int. Med. Suppl. 70:147-51], while P-glycoprotein expression was shown to serve as a molecular marker for

10   response to chemotherapy in bone and soft tissue sarcomas [Stein V et al. (1996) Eur. J. Cancer 32A:86-92].

Further progress towards understanding the clinical importance of MDR associated proteins and P-glycoprotein specifically, is hampered by the lack of validation methods to determine their expression.

15   *EXAMPLE 1*

Construction of a kit suitable for identifying and quantifying multi-drug

resistance associated proteins in a biological sample

The large variety of proteins, which are found to be associated with a condition or a disease, defines a need for a kit which can rapidly evaluate

protein antigen in a tissue sample, such as in a biopsy taken from a suspected cancer patient. The kit presented hereinunder is suitable for reacting multiple antigens present in a biological sample with a large number of antibodies at once.

5   *Step 1 - In-silico extraction of tryptic amino acid sequences:* Entire protein sequences of P-glycoprotein (GenBank accession number Hs.21330) and Mitoxantrone resistance protein (MXR) (GenBank accession number Hs. 194720), two of the numerous MDR associated proteins, were retrieved from the protein gene bank. The amino acid sequences of P-glycoprotein and MXR

10 were computationally analyzed to obtain tryptic amino acid sequences using the edit/replace function of Microsoft Word (Microsoft Incorporation). Tryptic sequences derived from P-glycoprotein are presented in Table 1 below.

Table 1

P-glycoprotein tryptic fragments	MXR tryptic fragments
MDLEGDR (SEQ ID NO: 1)	MSSSNVEVFIPVSQGNTNGFPATVSNDLK (SEQ ID NO: 150)
NGGAK (SEQ ID NO 2)	AFTEGAVLSFHNICYR (SEQ ID NO: 151)
K (SEQ ID NO: 3)	VK (SEQ ID NO: 152)
K (SEQ ID NO: 4)	LK (SEQ ID NO: 153)
NFFK (SEQ ID NO:5)	SGFLPCR (SEQ ID NO: 154)
LNNK (SEQ ID NO:6)	K (SEQ ID NO: 155)
SEK (SEQ ID NO 7)	PVEK (SEQ ID NO: 156)
DK (SEQ ID NO 8)	EILSNINGIMK (SEQ ID NO: 157)
K (SEQ ID NO: 9)	PGLNAILGPTGGGK (SEQ ID NO: 158)
EK (SEQ ID NO: 10)	SSLLDVLAAR (SEQ ID NO: 159)
K (SEQ ID NO: 11)	K (SEQ ID NO: 160)
PTVSVFSMFR (SEQ ID NO:12)	DPSGLSGDVLINGAPR (SEQ ID NO: 161)
YSNWLDK (SEQ ID NO:13)	PANFK (SEQ ID NO: 162)
LYMVVGTLAAIIHGAGLPLMMLVFGEMT DIFANAGNLEDLMSNITNR (SEQ ID NO:14)	CNSGYVVQDDVVMGTLTVR (SEQ ID NO: 163)

SDINDTGFFMNLEEDMTR (SEQ ID NO:15)	ENLQFSAALR (SEQ ID NO: 164)
YAYYYSGIGAGVLVAAYIQVSFWCLAAG R (SEQ ID NO: 16)	LATTMTNHEK (SEQ ID NO: 165)
QIHK (SEQ ID NO17)	NER (SEQ ID NO: 166)
IR (SEQ ID NO 18)	INR (SEQ ID NO: 167)
K (SEQ ID NO 19)	VIEELGLDK (SEQ ID NO: 168)
QFFHAIMR (SEQ ID NO 20)	VADSK (SEQ ID NO: 169)
QEIGWFDVHDVGELNTR (SEQ ID NO 21)	VGTQFIR (SEQ ID NO: 170)
LTDDVSK (SEQ ID NO 22)	GVSGGER (SEQ ID NO: 171)
INEGIGDK (SEQ ID NO 23)	K (SEQ ID NO: 172)
IGMFFQSMATFFTGFIVGFTR (SEQ ID NO: 24)	R (SEQ ID NO: 173)
GWK (SEQ ID NO 25)	TSIGMELITDPSILSLDEPTTGLDSSTANAV LLLLK (SEQ ID NO: 174)
LTLVILAISPVLGLSAAVWAK (SEQ ID NO:26)	R (SEQ ID NO: 175)
ILSSFTDK (SEQ ID NO: 27)	MSK (SEQ ID NO: 176)
ELLAYAK (SEQ ID NO: 28)	QGR (SEQ ID NO: 177)
AGAVAEEVLAAIR (SEQ ID NO: 29)	TIIFSIHQPR (SEQ ID NO: 178)
TVIAFGGQK (SEQ ID NO: 30)	YSIFK (SEQ ID NO: 179)
K (SEQ ID NO: 31)	LFDSLTLLASGR (SEQ ID NO: 180)
ELER (SEQ ID NO: 32)	LMFHGPAQEALGYFESAGYHCEAYNNPA DFFLDIINGDSTAVALNR (SEQ ID NO: 181)
YNK (SEQ ID NO: 33)	EEDFK (SEQ ID NO: 182)
NLEEAK (SEQ ID NO: 34)	ATEIIEPSK (SEQ ID NO: 183)
R (SEQ ID NO: 35)	QDK (SEQ ID NO: 184)
IGIK (SEQ ID NO: 36)	PLIEK (SEQ ID NO: 185)
K (SEQ ID NO: 37)	LAEIYVNSSFYK (SEQ ID NO: 186)
AITANISIGAAFLLIYASYALAFWYGTTLV LSGEYSIGQVLTVFFSVLIGAFSVGQASPSI EAFANAR (SEQ ID NO: 38)	ETK (SEQ ID NO: 187)
GAAYEIFK (SEQ ID NO: 39)	AELHQLSGGEK (SEQ ID NO: 188)
IIDNK (SEQ ID NO: 40)	K (SEQ ID NO: 189)
PSIDSYSK (SEQ ID NO: 41)	K (SEQ ID NO: 190)
SGHK (SEQ ID NO: 42)	K (SEQ ID NO: 191)
PDNIK (SEQ ID NO: 43)	ITVFK (SEQ ID NO: 192)
GNLEFR (SEQ ID NO: 44)	EISYTTSFCHQLR (SEQ ID NO: 193)
NVHFSYPSR (SEQ ID NO: 45)	WVSK (SEQ ID NO: 194)
K (SEQ ID NO: 46)	R (SEQ ID NO: 195)
EVK (SEQ ID NO: 47)	SFK (SEQ ID NO: 196)
ILK (SEQ ID NO: 48)	NLLGNPQASIAQIIVTVVLGLVIGAIYFGL K (SEQ ID NO: 197)
GLNLK (SEQ ID NO: 49)	NDSTGIQNR (SEQ ID NO: 198)
VQSGQTVALVGNSGCGK (SEQ ID NO: 50)	AGVLFFLTTNQCFSSVSAVELFVVEK (SEQ ID NO: 199)
STTVQLMQR (SEQ ID NO: 51)	K (SEQ ID NO: 200)
LYDPTEGMVSVDGQDIR (SEQ ID NO: 52)	LFIHEYISGYYR (SEQ ID NO: 201)
TINVR (SEQ ID NO: 53)	VSSYFLGK (SEQ ID NO: 202)
FLR (SEQ ID NO: 54)	LLSDLLPMR (SEQ ID NO: 203)
EIIGVVSQEPVLFATTIAENIR (SEQ ID NO: 55)	MLPSIIFTCIVYFMLGLK (SEQ ID NO: 204)
YGR (SEQ ID NO: 56)	PK (SEQ ID NO: 205)
ENVTMDEIEK (SEQ ID NO: 57)	ADAFFVMMFTLMMVAYSASSMALAIAA GQSVVSVATLLMTICFVFMMIFSGLLVNL TTIASWLSWLQYFSIPR (SEQ ID NO: 206)
AVK (SEQ ID NO: 58)	YGFTALQHNEFLGQNFCPGLNATGNNPC NYATCTGEEYLVK (SEQ ID NO: 207)

EANAYDFIMK (SEQ ID NO: 59)	QGIDLSPWGLWK (SEQ ID NO: 208)
LPHK (SEQ ID NO: 60)	NHVALACMIVIFLTIAYLK (SEQ ID NO: 209)
FDTLVGER (SEQ ID NO: 61)	LLFLK (SEQ ID NO: 210)
GAQLSGGQK (SEQ ID NO: 62)	K (SEQ ID NO: 211)
QR (SEQ ID NO: 63)	YS (SEQ ID NO: 212)
IAIAR (SEQ ID NO: 64)	
ALVR (SEQ ID NO: 65)	
NPK (SEQ ID NO: 66)	
ILLLDEATSALDTESEAVVQVALDK (SEQ ID NO: 67)	
AR (SEQ ID NO: 68)	
K (SEQ ID NO: 69)	
GR (SEQ ID NO: 70)	
TTIVIAHR (SEQ ID NO: 71)	
LSTVR (SEQ ID NO: 72)	
NADVIAGFDDGVIVEK (SEQ ID NO: 73)	
GNHDELMK (SEQ ID NO: 74)	
EK (SEQ ID NO: 75)	
GIYFK (SEQ ID NO: 76)	
LVTMQTAGNEVELENAADESK (SEQ ID NO: 77)	
SEIDALEMSSNDSR (SEQ ID NO: 78)	
SSLIR (SEQ ID NO: 79)	
K (SEQ ID NO: 80)	
R (SEQ ID NO: 81)	
STR (SEQ ID NO: 82)	
R (SEQ ID NO: 83)	
SVR (SEQ ID NO: 84)	
GSQAQDR (SEQ ID NO: 85)	
K (SEQ ID NO: 86)	
LSTK (SEQ ID NO: 87)	
EALDESIPPVSFWR (SEQ ID NO: 88)	
IMK (SEQ ID NO: 89)	
LNLTEWPYFVVGVFCAIINGGLQPAFAIIFSK (SEQ ID NO: 90)	
IIGVFTR (SEQ ID NO: 91)	
IDDPETK (SEQ ID NO: 92)	
R (SEQ ID NO: 93)	
QNSNLFSLLFLALGIISFITFFLQGFTFGK AGEILTK (SEQ ID NO: 94)	
R (SEQ ID NO: 95)	
LR (SEQ ID NO: 96)	
YMVFR (SEQ ID NO: 97)	
SMLR (SEQ ID NO: 98)	
QDVSWFDDPK (SEQ ID NO: 99)	
NTTGALTTR (SEQ ID NO: 100)	
LANDAAQVK (SEQ ID NO: 101)	
GAIGSR (SEQ ID NO: 102)	
LAVITQNIANLGTGIISFIYGWQLTLLLLAIVPIIAIAGVVEMK (SEQ ID NO: 103)	
MLSGQALK (SEQ ID NO: 104)	
DK (SEQ ID NO: 105)	
K (SEQ ID NO: 106)	
ELEGAGK (SEQ ID NO: 107)	
IATEAIENFR (SEQ ID NO: 108)	
TVVSLTQEQK (SEQ ID NO: 109)	

FEHMYAQSLQVPYR (SEQ ID NO: 110)
NSLR (SEQ ID NO: 111)
K (SEQ ID NO: 112)
AHIFGITFSFTQAMMYFSYAGCFR (SEQ ID NO: 113)
FGAYLVAHK (SEQ ID NO: 114)
LMSFEDVLLVFSAVVFGAMAVGQVSSFA PDYAK (SEQ ID NO: 115)
AK (SEQ ID NO: 116)
ISAAHIIMIIEK (SEQ ID NO: 117)
TPLIDSYSTEGLMPNTLEGNVTFGEVVFN YPTR (SEQ ID NO: 118)
PDIPVLQGLSLEVK (SEQ ID NO: 119)
K (SEQ ID NO: 120)
GQTLALVGSSGCGK (SEQ ID NO: 121)
STVVQLLER (SEQ ID NO: 122)
FYDPLAGK (SEQ ID NO: 123)
VLLDGK (SEQ ID NO: 124)
EIK (SEQ ID NO: 125)
R (SEQ ID NO: 126)
LNVQWLR (SEQ ID NO: 127)
AHLGIVSQEPILFDCSIAENIAYGDNSR (SEQ ID NO: 128)
VVSQEEIVR (SEQ ID NO: 129)
AAK (SEQ ID NO: 130)
EANIHAFIESLPNK (SEQ ID NO: 131)
YSTK (SEQ ID NO: 132)
VGDK (SEQ ID NO: 133)
GTQLSGGQK (SEQ ID NO: 134)
QR (SEQ ID NO: 135)
IAIAR (SEQ ID NO: 136)
IAIAR (SEQ ID NO: 137)
ALVR (SEQ ID NO: 138)
QPHILLLDEATSALDTESEK (SEQ ID NO: 139)
VVQEALDK (SEQ ID NO: 140)
AR (SEQ ID NO: 141)
EGR (SEQ ID NO: 142)
TCIVIAHR (SEQ ID NO: 143)
LSTIQNADLIVVFQNGR (SEQ ID NO: 144)
VK (SEQ ID NO: 145)
EHGTHQQLLAQK (SEQ ID NO: 146)
GIYFSMVSVQAGTK (SEQ ID NO: 147)
R (SEQ ID NO: 148)
Q (SEQ ID NO: 149)

Table 1 Cont.

Step 2 - In silico selection of non-homologous amino acid sequences:

Computationally extracted tryptic amino acid sequences of P-glycoprotein and

MXR protein were scanned for homology to all known protein sequences using

the BLAST and Smith-Waterman algorithms by Unix-interfaced GCG server.

Only a portion of the tryptic amino acid sequences listed in Table 1,

were found to be unique to each of P-glycoprotein and MXR protein. These

5   sequences, which were not found in any other human protein recorded in the

human database (Unigene databank) are listed in Table 2 below.

Table 2

	P-glycoprotein Unique tryptic fragments	MXR Unique tryptic fragments
1.	LYMVVGTLAAIIHGAGLPLMMLVF GEMTDIFANAGNLEDLMSNITNR (SEQ ID NO: 213)	MSSSNVEVFIPVSQGNTNGFPATVSNDLK (SEQ ID NO: 234)
2.	SDINDTGFFMNLEEDMTR (SEQ ID NO: 214)	AFTEGAVLSFHNICYR (SEQ ID NO: 235)
3.	YAYYYSGIGAGVLVAAYIQVSFWCL AAGR (SEQ ID NO: 215)	EILSNINGIMKPGLNAILGPTGGGK (SEQ ID NO: 236)
4.	IGMFFQSMATFFTGFIVGFTR (SEQ ID NO: 216)	DPSGLSGDVLINGAPRPANFK (SEQ ID NO: 237)
5.	LTLVILAISPVLGLSAAVWAK (SEQ ID NO: 217)	CNSGYVVQDDVVMGTLTVR (SEQ ID NO: 238)
6.	AITANISIGAAFLLIYASYALAFWYG TTLVLSGEYSIGQVLTVF FSVLIGAFSVGQASPSIEAFANAR (SEQ ID NO: 218)	ENLQFSAALR (SEQ ID NO: 239)
7.	LYDPTEGMVSVDGQDIR (SEQ ID NO: 219)	LATTMTNHEK (SEQ ID NO: 240)
8.	ILLLDEATSALDTESEAVVQVALDK (SEQ ID NO: 220)	TSIGMELITDPSILSLDEPTTGLDSSTANA VLLLLK (SEQ ID NO: 241)
9.	NADVIAGFDDGVIVEK (SEQ ID NO: 221)	TIIFSIHQPR (SEQ ID NO: 242)
10.	LVTMQTAGNEVELENAADESK (SEQ ID NO: 222)	LFDSLTLLASGR (SEQ ID NO: 243)
11.	SEIDALEMSSNDSR (SEQ ID NO: 223)	LMFHGPAQEALGYFESAGYHCEAYNNPA DFFLDIINGDSTAVALNR (SEQ ID NO: 244)
12.	EALDESIPPVSFWR (SEQ ID NO: 224)	LAEIYVNSSFYK (SEQ ID NO: 245)
13.	LNLTEWPYFVVGVFCAIINGGLQPA FAIIFSK (SEQ ID NO. 225)	EISYTTSFCHQLR (SEQ ID NO: 246)
14.	QNSNLFSLLFLALGIISFITFFLQGFTK (SEQ ID NO: 226)	NLLGNPQASIAQIIVTVVLGLVIGAIYFGLK (SEQ ID NO: 247)

15.	LAVITQNIANLGTGIIISFIYGWQLTL LLLAIVPIIAIAGVVEMK (SEQ ID NO: 227)	AGVLFFLTTNQCFSSVSAVELFVVEK (SEQ ID NO: 248)
16.	FEHMYAQSLQVPYR (SEQ ID NO: 228)	LFIHEYISGYYR (SEQ ID NO: 249)
17.	AHIFGITFSFTQAMMYFSYAGCFR (SEQ ID NO: 229)	ADAFFVMMFTLMMVAYSASSMALAIAAGQSV SVATLLMTICFVFMMIFSGLLVNLTTIASWLSWL QYFSIPR (SEQ ID NO: 250)
18.	LMSFEDVLLVFSAVVFGAMAVGQS SFAPDYAK (SEQ ID NO: 230)	YGFTALQHNEFLGQNFCPGLNATGNNPCN YATCTGEEYLVK (SEQ ID NO: 251)
19.	TPLIDSYSTEGLMPNTLEGNVTFGEV VFNYPTR (SEQ ID NO: 231)	QGIDLSPWGLWK (SEQ ID NO: 252)
20.	EANIHAFIESLPNK (SEQ ID NO: 232)	NHVALACMIVIFLTIAYLK (SEQ ID NO: 253)
21.	GIYFSMVSVQAGTK (SEQ ID NO: 233)	

Table 2 Cont.

Step 3 - In silico selection of immunogenic amino acid sequences:

These unique tryptic amino acid sequences were further analyzed for immunogenicity, by testing parameters such as foreignness, amino acid chemical composition and heterogeneity, peptide molecular weight and susceptibility to antigen processing and presentation.

Following such analysis, several immunogenic sequence candidates were selected from the amino acid sequences presented in Table 2. The candidate sequences selected included peptides: 2, 7, 8, 11 and 19 of the P-glycoprotein list and peptides 5, 8, 10, 13 and 16 of the MXR list.

Step 4 - Preparation of selected amino acid sequences: Peptide 8 of

the MXR list and peptides 8 and 19 of the P-glycoprotein list were selected from the above described candidates for further studies. Highly immunogenic

portions of these peptides including amino acids 9-21 of MXR peptide 8 and

amino acids 12-25 and 13-22 of P-glycoprotein peptides 8 and 19

(respectively) were synthesized and purified by HPLC.

Step 5 - Generation of antibodies against selected peptides: Polyclonal

5 antibodies directed against the selected amino acid sequence of MXR were

prepared in rabbit using known techniques, while monoclonal antibodies were

obtained against the selected amino acid sequences of the P-glycoprotein.

Step 6 - Matrix construction and calibration: To calibrate the matrix,

as to non-specifically bound proteins, polyclonal antibodies against three

10 proteins known to be constitutively expressed in human tissues (i.e., house

keeping proteins) were further prepared in rabbits. Each of the control

antibodies and the antibodies generated in step 4 above, were conjugated to a

solid support, comprised of a multiwell plastic plate (Nunc Immunosorb).

Antibodies were left to bind overnight at 4°C at a neutral pH in

15 phosphate-buffered saline (PBS). Subsequent to antibody conjugation, the

multiwell plate was washed in PBS containing 1% bovine serum albumine

(BSA), and finally stored at 4°C in the presence of 0.01% sodium azide thus

generating a matrix of substrate-bound antibodies each specifically directed

against an MDR associated peptide.

Step 7 - preparation of purified labeled tryptic peptides capable of binding the matrix antibodies: All tryptic peptides used for generating the antibodies of step 5 were end labeled at the amino terminus by fluorescent labeling, such that a uniform signal is obtained in all the wells or positions in the matrix when these peptides are conjugated with their respective antibodies. This mixture of labeled tryptic polypeptides is capable of competing with MDR associated proteins found in a sample of interest.

EXAMPLE 2

A competition assay for quantifying multi-drug resistance associated proteins in a biological sample

This example illustrates a stepwise procedure, which utilizes the matrix described above to identify and quantify multi-drug resistance associated proteins in a biological sample.

A cell culture, which exhibits an MDR phenotype, is dissolved by suspending the sample in a 1-2 % sodium dodecyl sulfate (SDS) containing buffer for one hour. Denatured proteins are precipitated by adding methanol/acetic acid (pH-4) followed by an overnight incubation at -20°C. Thereafter, the protein precipitate is resuspended in 0.05-0.1 % SDS, and trypsin is added to the resuspended sample. Tryptic digestion is allowed to

proceed for 16 hours at 37° until complete proteolytic fragmentation of the sample proteins. At the completion of digestion residual tryptic activity is terminated by adding bovine trypsin inhibitor.

Subsequently, a portion of the tryptic-digested sample is added to an antibody matrix, which is prepared as described hereinabove. Incubation is allowed to proceed for 1 hour at room temperature to allow formation of immunocomplexes, following which, the matrix is washed twice with phosphate buffered saline (PBS) containing 1 % BSA, to reduce non-specific binding.

Monitoring specific binding of sample proteins to the matrix is effected using a mixture of fluorescently labeled polypeptides against which the matrix antibodies were raised. In principle, such a peptide mixture is designed to generate a homogeneously fluorescent surface when applied on a fresh matrix, and to provide a reduced signal when applied to a matrix, which has been previously treated with a digested protein sample detectable by the matrix. The intensity of the fluorescent signal obtained from the matrix as correlated to specific positions in the matrix gives a quantitative measure of the amount of protein present in the sample, after considering the amount of protein sample

applied to the matrix, and control binding, as determined using signals obtained from three control antibodies described in Example 1.

EXAMPLE 3

A direct assay for quantifying multi-drug resistance associated proteins in a

5
biological sample

This assay is aimed at quantifying MDR associated proteins in a biological sample by utilizing fluorescently labeled peptides products of the digested sample. As such, this assay is based on a direct correlation between the level of the fluorescent signal obtained and the level of the protein to be

10   quantified.

A labeled digested sample is applied onto a fresh antibody matrix, and specific immunocomplexes are allowed to form as described in Example 2 of the Examples section. The intensity of fluorescent signal correlated with specific positions in the matrix, gives a quantitative measure of the amount of

15   MDR associated proteins present in the sample, after considering the amount of protein sample applied to the matrix, and control binding, as determined using signals obtained from three control antibodies described in Example 1.

EXAMPLE 4

Quantifying P-glycoprotein expression level in membrane fractions of

chinese hamster ovarian cell-lines

The following Example illustrates use of an MDR1-specifc matrix, for

5   determining the expression level of P-glycoprotein (i.e., MDR1 gene product)

in two different Chinese hamster ovarian cell-lines.

Method: Total membrane protein (1700 μg) obtained from

P-glycoprotein expressing CHO cells (Pgp-CHO) and control wild type cells

(WT-CHO) was suspended in 1.12 % SDS, precipitated with methanol/acetic

10   acid, incubated at -20°C for 16 hours, resuspended in 0.07 % SDS and digested

overnight with N-tosyl-L-phenylalanine chloromethylketone (TPCK)-treated

trypsin.   Digested protein samples were centrifuged at 14,000 rpm for 25

minutes, and then spin-filtered through a Vivaspin column (Vivascience Ltd,

UK), to remove undigested sample.

15   The digest was then applied onto each well of an MDR1 specific matrix

which was precoated with a monoclonal antibody generated against the peptide

probe, MPNTLEGNVTK, where all but the C-terminal lysine corresponded to

the central section of an informatically derived tryptic digest product of

P-glycoprotein (amino acid coordinates 13-22 of SEQ ID NO: 231).   Three

control wells of the 96 well matrix were pre-coated with 250 μg of C494, a commercially available monoclonal antibody recognizing P-glycoprotein (Dako corporation, USA).

Serial dilutions of the peptide probe were added in duplicates along with

5 the digested samples to each well starting at minimal dilution containing 500 ng of probe peptide. The matrix was further supplemented with 10 ng of biotinylated probe peptide, and the resulting mixture was incubated for 2 hr room temperature.

Following incubation, the matrix was washed with a washing buffer

10 containing 1 % BSA, and subsequently reacted with streptavidin conjugated to horse-radish peroxidase. The obtained color reaction of the peroxidase enzyme with substrate TMB 3,3',5,5' – Tetramethylbenzidine measured the amount of biotin present in each well. Optical density was determined using an Elisa reader.

15 ***Results:*** As shown in Figure 3, digested membrane sample of wild-type CHO cells exhibited no reduction in optical density when compared to a control sample, which lacked the labeled digest products, suggesting that wild-type CHO cells are devoid of P-glycoprotein expression. In contrast, digested membrane samples of P-glycoprotein containing cells (Pgp-CHO),

exhibited a significant decrease in optical density. As was extrapolated from the calibration curve, 500 µg of original membrane protein represents 0.025 µg of peptide. Given that the molecular weight of the probe peptide is 1000 Da, and that P-glycoprotein has a total molecular weight of 14,460 Da, the content

5   of P-glycoprotein in each well corresponded to 9.33 µg which represents 1.64 % of the cell membrane protein mass. Turnover analysis and enzymatic activity of P-glycoprotein in CHO membranes suggests that the P-glycoprotein membrane content was 5.6 %, implying that the quantitative analysis recovered 29 % of the enzyme originally present.

10   Control experiments in which a probe peptide was added to the membrane sample prior to the digestive step and subsequently applied on the MDR1-matrix showed an expected recovery of 45 %. The similarity of these two values further validates the accuracy and efficiency of the matrix constructed according to the teachings of the present invention in identifying

15   and quantifying proteins in a biological sample.

EXAMPLE 5

A stepwise usage of an antibody matrix to identify and quantify levels of a

disease associated protein

Figure 4 illustrate a method of identifying and quantifying disease

5   associated proteins using a protein-specific antibody matrix kit designed and

constructed according to the teachings of the present invention.

Step 1: a biopsy (1), which is taken from a patient is solubilized and

incubated with a proteolytic agent until complete protein cleavage is achieved

(2).

10   *Step 2:* The digested sample is mixed with a sample of tagged peptides

(3) (represented in the figure by arrow-attached hexagons). The labeled

peptides are specific for one or more proteins of interest. These peptides have

been selected by scanning a protein database of human sequences for peptides

that are unique to protein(s) of interest. Circles, cylinders, non-arrowed

15   hexagons and other shapes (4) represent the peptides in the biopsy digest.

Step 3: The mixture is layered (7) onto a matrix (5), which includes a

set of antibodies prepared against the selected peptides. The matrix can be a

multiwell plate (e.g., 96 wells) onto which the antibodies are directly or

indirectly attached in a regiospecific manner. The inset illustrates the

antibodies attached to the wells (6); two wells are depicted, each having a particular antibody attached to it. In addition, the matrix may also include wells of control antibodies, which are prepared against proteins that are ubiquitously present in tissue samples. Furthermore, additional wells may also

5   include labeled peptides at known amounts to serve as standards from which a calibration curve can be derived.

Step 4: The mixture of selected peptides, tagged peptides and attached antibodies(8) is left to form immunocomplexes (9). Inset (10) depicts formed immunocomplexes.

10   *Step 5:* The matrix is washed several times with a blocking buffer (11) in order to free non-specifically attached peptides.

Step 6: An enzyme-linked agent specific for the tagged peptides is added (12) and the wells are washed again in order to remove non-specifically bound agent.

15   *Step 7:* The matrix is incubated (13) with a substrate, which generates a color reaction (14) when processed by the enzyme-linked agent described in Step 6.

Step 8: The intensity of the color reaction (15) produced is measured in a conventional 96-well plate reader and the data analyzed by a computer

program(16) which determines the amount of peptide present in each of the wells.

Although the invention has been described in conjunction with specific

5   embodiments thereof, it is evident that many alternatives, modifications and variations will be apparent to those skilled in the art. Accordingly, it is intended to embrace all such alternatives, modifications and variations that fall within the spirit and broad scope of the appended claims. All publications, patents, patent applications and sequences identified by their accession

10   numbers mentioned in this specification are herein incorporated in their entirety by reference into the specification, to the same extent as if each individual publication, patent, patent application or sequence identified by their accession number was specifically and individually indicated to be incorporated herein by reference. In addition, citation or identification of any

15   reference in this application shall not be construed as an admission that such reference is available as prior art to the present invention.